# On Projections of Metric Spaces

Mark Kozdoba

**Abstract**

Let $X$ be a metric space and let $\mu$ be a probability measure on it. Consider a Lipschitz map $T : X \to \mathbb{R}^n$, with Lipschitz constant $\leq 1$. Then one can ask whether the image $TX$ can have large projections on many directions. For a large class of spaces $X$, we show that there are directions $\phi \in S^{n-1}$ on which the projection of the image $TX$ is small on the average, with bounds depending on the dimension $n$ and the eigenvalues of the Laplacian on $X$.

## 1  Introduction

Let $(X, m)$ be a metric space and let $\mu$ be a probability measure on $X$. Consider a Lipschitz map $T : X \to \mathbb{R}^n$, with $\|T\|_{Lip} \leq 1$, where $\mathbb{R}^n$ is taken with the standard inner product $< \cdot, \cdot >$ and the corresponding Euclidean norm $|\cdot|$.

Define another (semi) norm on $\mathbb{R}^n$ by setting for each $\theta \in \mathbb{R}^n$,

$$\|\theta\|_{L_2} := \left( \int < Tx, \theta >^2 d\mu(x) \right)^{\frac{1}{2}}.$$

This norm is known as the covariance structure of the push-forward measure $T\mu$, and we shall regard it as measuring the size of the projection of the image of $X$ onto the direction $\theta$. It is then natural to ask, for a given space $X$, whether an image of $X$ can have big projections on many directions. For example, take $X$ to be the unit interval $[0, 1]$, with Lebesgue measure, and let $T : [0, 1] \to \mathbb{R}^n$ be a Lipschitz map (here and in what follows, Lipschitz will mean "having Lipschitz constant $\leq 1$"). By considering a few pictures, it is clear that if $n$ is large then there should be a direction with a small projection and we are interested in quantifying this phenomenon. It turns out that this particular situation was investigated already by Gohberg and Krein, in a different formulation and context (related to spectral properties of smooth kernels), see the book [GK], Chapter III.10.3 .

**Theorem 1.1** ( [GK]). *There is an absolute constant $c > 0$, such that for every $n > 1$ and every $T : [0, 1] \to \mathbb{R}^n$ with $\|T\|_{Lip} \leq 1$, there is $\theta \in S^{n-1}$ such that $\|T\|_{L_2} \leq c \cdot n^{-\frac{3}{2}}$.*

Our objective is to provide a similar result for general metric spaces with measure on which an appropriate notion of a derivative can be defined. However, for the simplicity of presentation, we shall restrict our discussion to finite spaces $X$ which are graphs with the shortest path metric. The statement and the proof of the following result can, for instance, can be repeated with straightforward modifications in the setting of Riemannian manifolds.

Let $X$ be a finite set and let $E \subset X \times X$ be a set of edges such that $(X, E)$ is a connected undirected graph without loops. For $x, y \in X$, write $x \sim y$ iff $(x, y) \in E$ and let $d(x)$ denote the degree of $x$. We endow $X$ with the shortest path metric and we set $\mu$ to be the stationary distribution of a simple nearest neighbour random walk on $(X, E)$, $\mu(x) = \frac{d(x)}{2|E|}$. Denote by $L(X)$ the space of real-valued functions on $X$. The Laplacian on $L(X)$ is defined by

$$\triangle(f)(x) = 2 \left( f(x) - \frac{1}{deg(x)} \sum_{y \sim x} f(y) \right).$$

As is well known, this is a self-adjoint non-negative operator, with a one dimensional kernel consisting of constant functions. We denote by $\{\lambda_i\}_{i=1}^{|X|-1}$ the sequence of non-zero eigenvalues of $\triangle$ in *non-decreasing* order, including multiplicities.

**Theorem 1.2.** *There is an absolute constant $c > 0$ such that for every graph $X$ as above, every $n > 1$ and every Lipschitz map $T : X \to \mathbb{R}^n$, there is a direction $\theta \in S^{n-1}$ such that $\|\theta\|_{L_2} \leq c \cdot n^{-\frac{1}{2}} \lambda_{\lfloor n/2 \rfloor -1}^{-\frac{1}{2}}$.*

*Example* 1.3 (Discrete Space). Let $(X, m)$ be a metric space such that $m(x, y) = 1$ for all $x \neq y$, and let $\mu$ be the uniform probability on $X$. Then for every Lipschitz $T : X \to \mathbb{R}^n$ there is $\theta \in \mathbb{R}^n$ such that

$$\|\theta\|_{L_2} \leq \frac{c}{\sqrt{n}}. \tag{1}$$

Note that, interestingly enough, the size of the space, $|X|$, does not appear in this bound (except that, of course, for $n > |X|$ we always have $\theta \in \mathbb{R}^n$ with $\|\theta\|_{L_2} = 0$). That is, one can not increase the minimal projection size of $X$ in, say, $\mathbb{R}^{20}$, by adding more points. To prove (1), note that $(X, m)$ corresponds to a clique graph, and the Laplacian has a single non-zero eigenvalue, which is of the order of a constant and has multiplicity $|X| - 1$.

*Example* 1.4 (Combinatorial Cube). Here $X$ is the set $\mathcal{C}^d = \{0,1\}^d$, with the Hamming metric

$$m(x,y) = |\{i \in \{1,...,d\} \mid x_i \neq y_i\}|$$

where $x = x_1...x_d, y = y_1...y_d \in \mathcal{C}^d$ and $\mu$ is again taken to be the uniform probability measure. The non-zero eigenvalues of the Laplacian on $X$ are $\frac{4k}{d}$ with multiplicity $\binom{n}{k}$, $k = 1,...,d$ (See, ex. [Diac]). The $n$-th smallest eigenvalue is therefore of the order $\frac{\log(n)}{d}$ and we obtain that for every Lipschitz $T : X \to \mathbb{R}^n$, there is $\theta \in S^{n-1}$ such that

$$\|\theta\|_{L_2} \leq c \cdot \frac{\sqrt{d}}{\sqrt{n} \cdot \sqrt{\log(n)}}$$

Our approach to Theorem 1.2 extends the argument in [GK] and, naturally, Theorem 1.1 can also be seen as a consequence of the principle of Theorem 1.2. Indeed, the eigenvalues of the Laplacian on $[0,1]$ satisfy, up to multiplicative constants, that $\lambda_n \approx n^2$ (as can be seen by double differentiating sines and cosines). Hence the decay rate of $n^{-\frac{1}{2}} \cdot (n^2)^{-\frac{1}{2}} = n^{-\frac{3}{2}}$ in Theorem 1.1.

# 2   Proof

Fix a graph $(X, E)$ with the stationary measure $\mu$ on it. Let $T : X \to \mathbb{R}^n$ be a Lipschitz map. It will be convenient to introduce an additional assumption, that

$$\int T(x)d\mu = 0. \tag{2}$$

With conditions of Theorem 1.2 and this assumption, we will show that there is $\theta \in S^{n-1}$ such that

$$\|\theta\|_{L_2} = c \cdot n^{-\frac{1}{2}} \lambda_{\lfloor n/2 \rfloor}^{-\frac{1}{2}}. \tag{3}$$

This implies the result for arbitrary Lipschitz $T : X \to \mathbb{R}^n$. Indeed, assume that $v := \int Tx d\mu \neq 0$. Denote by $P$ the orthogonal projection onto the $n-1$ dimensional space orthogonal to $v$. Then the composition $P \circ T$ is a Lipschitz map into that space and $\int (P \circ T)(x)d\mu = 0$, so (3) can be applied in dimension $n-1$.

The notion of a gradient on a graph is folklore, although not particularly frequent in the literature. We thus recall the definition.

Define an inner product on $L(X)$ by

$$[f, g]_X = \int f(x)g(x)d\mu(x).$$

Denote by $L(E)$ the set of real valued functions on the set of edges, $E$, and let $\nu$ be the uniform probability measure on $E$. We equip $L(E)$ with the inner product

$$[f, g]_E = \int f(e)g(e)d\nu.$$

Fix an arbitrary orientation on $E$, i.e. for each edge $e = \{x, y\}$ choose in an arbitrary way an enumeration $v_e^1, v_e^2$ of the vertices of the edge. The gradient operator on $X$ is defined by $\bigtriangledown : L(X) \to L(E)$,

$$(\bigtriangledown f)(e) = f(v_e^1) - f(v_e^2).$$

One can verify by direct computation that the usual relation $\triangle = \bigtriangledown^* \circ \bigtriangledown$ holds, where $\bigtriangledown^*$ is the Hilbert space adjoint of $\bigtriangledown$. Note that the Laplacian does not depend on the orientation on $E$.

Denote by $L_0(X)$ the subspace of $L(X)$ that is orthogonal to the constant functions. Since $Ker\triangle$ is the space of the constant functions, the Laplacian is invertible on $L_0(X)$ and hence the gradient is also an invertible operator from $L_0(X)$ onto its image. The inverse of the gradient will be of importance in what follows and we denote it by $\Gamma : Im(\bigtriangledown) \to L_0(X)$.

We also recall the notion of singular values of an operator. If $V$ and $W$ are (say, finite dimensional ) Hilbert spaces, and $A : V \to W$ is a linear operator, then $A^*A$ is a non-negative self-adjoint operator. Let $\{\lambda_i(A^*A)\}_{i=1}^{dim V}$ be the eigenvalues of $A^*A$, in non-increasing order, with multiplicities. Then singular values of the operator $A$ are the non-increasing sequence $s_i(A) = \sqrt{\lambda_i(A^*A)}$.

The singular values of $A$ have a simple geometric interpretation. If $B_2(V)$ denotes the unit ball of $V$, then the image $A(B_2(V))$ is an ellipsoid in $W$ and $s_i(A)$ are precisely the lengths of the principal axes of this ellipsoid.

The singular values satisfy $s_i(A) = s_i(A^*)$ and singular values of a composition can be bounded by the following inequality due to Ky Fan. Let $A : U \to W$ and $B : W \to V$ be two operators on Hilbert spaces. Then, for every $i, j \geq 1$,

$$s_{i+j-1}(BA) \leq s_i(A)s_j(B) \tag{4}$$

Finally, the Hilbert-Schmidt norm of an operator is defined by

$$\|A\|_{HS} = \sqrt{tr A^*A} = \left( \sum_i s_i^2(A) \right)^{\frac{1}{2}}.$$

4

Note that since $s_i$ is a non-increasing sequence,

$$s_i(A) \leq \frac{\|A\|_{HS}}{\sqrt{i}} \tag{5}$$

for all $i$.

More details on singular values can be found, for instance, in [Bha] or [GK].

*Proof of Theorem 1.2.* As mentioned above, we assume that condition (2) holds. Let $D_T$ denote the gradient of the map $T$, i.e.

$$D_T(e) = T(v_e^1) - T(v_e^2).$$

Since $T$ is Lipschitz, $D_T$ is bounded, i.e. $|D_T(e)| \leq 1$ for all $e \in E$.

For every $\theta \in \mathbb{R}^n$ consider the function on $X$, $f_\theta(x) = <\theta, Tx>$. By (2), $f_\theta \in L_0(E)$ and clearly $(\nabla f_\theta)(e) = <\theta, D_T(e)>$ and

$$f_\theta = \Gamma(<\theta, D_T>).$$

Let

$$B_{L_2}(X) = \left\{ f : X \to \mathbb{R} \;\middle|\; [f,f]_X \leq 1 \right\}$$

denote the unit ball in $L(X)$ and write $\|\theta\|_{L_2}$ in a dual form:

$$\|\theta\|_{L_2} = sup_{g \in B_2(X)} \int g(x) <\theta, Tx> d\mu.$$

Then

$$\int g(x) <\theta, Tx> d\mu = [g, f_\theta]_X = [g, \Gamma \circ \nabla f_\theta]_X = [\Gamma^* g, \nabla f_\theta]_E$$

Next, write

$$[\Gamma^* g, \nabla f_\theta]_E = \int (\Gamma^* g)(e) \cdot <\theta, D_T(e)> d\nu(e) = <\theta, \int (\Gamma^* g)(e) \cdot D_T(e) d\nu(e)>.$$

Denote by $\widetilde{D} : L(E) \to \mathbb{R}^n$ the operator that acts by

$$\widetilde{D}u = \int u(e) \cdot D_T(e) d\nu$$

.

With this notation,

$$\|\theta\|_{L_2} = sup_{g \in B_2} <\theta, \widetilde{D} \circ \Gamma^* g> = \sup_{\phi \in \mathcal{E}} <\theta, \phi> \tag{6}$$

5

where $\mathcal{E} = \widetilde{D} \circ \Gamma^* B_2$ is the dual (in $\mathbb{R}^n$) ellipsoid of the $\| \cdot \|_{L_2}$ norm.

Our aim is to bound the quantity

$$\inf_{\theta \in S^{n-1}} \|\theta\|_{L_2}.$$

By (6), this quantity equals the length of the smallest principal axe of the ellipsoid $\mathcal{E}$. Since the lengths of the principal axes are the singular values of $\widetilde{D} \circ \Gamma^*$, we bound the singular values of this operator.

The singular values of $\Gamma^*$ are given, $s_i(\Gamma^*) = \lambda_i^{-\frac{1}{2}}$, where $\lambda_i$ are the non-zero eigenvalues of the Laplacian (in non-decreasing order, so that $s_i(\Gamma^*)$ do not increase). To bound the singular values of $\widetilde{D}$, we show that

$$\|\widetilde{D}\|_{HS} \leq 1. \tag{7}$$

Indeed, one readily verifies that

$$\widetilde{D} \circ \widetilde{D}^*(\theta) = \int < \theta, D_T(e) > D_T(e) d\nu.$$

For fixed $e \in E$, the trace of an operator $\theta \mapsto < \theta, D_T(e) > D_T(e)$ is $|D_T(e)|^2 \leq 1$, implying (7). Now, by (5), $s_i(\widetilde{D}) \leq \frac{1}{\sqrt{i}}$ and an application of (4) (with $i = j = n/2$) completes the proof. $\square$

# References

[Bha] R. Bhatia, Matrix analysis. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997. xii+347

[Diac] P. Diaconis, Group representations in probability and statistics, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11, Institute of Mathematical Statistics, Hayward, CA, 1988, vi+198

[GK] I. C. Gohberg, M. G. Kreĭn , Introduction to the theory of linear non-selfadjoint operators, Translations of Mathematical Monographs, Vol. 18, American Mathematical Society,1969,xv+378